

10/ 039,369

IBM/196

In the Claims

1. (Currently Amended) A method for estimating statistics on an attribute of a relation, comprising

forming a histogram of said attribute of said relation, the histogram comprising statistics regarding tuples falling within each of a plurality of ranges of values of said attribute, said statistics being augmented to identify the most frequent values of an said attribute within at least one of said ranges,

evaluating said histogram in connection with a criterion for retrieval of data from a relation.

2. (Original) The method of claim 1 further comprising forming a second histogram of said attribute of a second relation, said second histogram being augmented to identify the most frequent values of said attribute, and

evaluating said histograms to identify frequent values shared by said histograms.

3. (Currently Amended) The method of claim 2, further comprising multiplying frequencies of frequent values in each of said histograms to produce a estimate of join fanout of a join of said relations on said attribute.

10/ 039,369

IBM/196

4. (Currently Amended) The method of claim 3, further comprising multiplying a number frequency of a frequent value in one said histogram by an estimate of the average number infrequent values in the other histogram.

5. (Original) The method of claim 3 further comprising computing a number of matching infrequent values in each said histogram by
estimating a number of infrequent values in each relation using said histograms, and
computing from said estimates the join fanout attributable to said attribute.

6. (Currently Amended) A computer system for implementing a relational database system and performing a user query on said relational database system, comprising
storage for relations of said relational database system, and a histogram of an attribute of a first of said relations, the histogram comprising statistics regarding tuples falling within each of a plurality of ranges of values of said attribute, said statistics being augmented to identify the most frequent values of ~~an~~ said attribute within at least one of said ranges in said first relation,
a computing circuit for implementing said relational database system, said computing circuit computing a statistic on said attribute by evaluating said histogram in connection with a criterion for retrieval of data from a relation.

10/ 039,369

IBM/196

7. (Original) The computer system of claim 6 wherein

said storage further includes a second histogram of said attribute of a second of said relations, said histogram being augmented to identify the most frequent values of said attribute in said second relation, and

said computing circuit evaluates said histograms to identify frequent values shared by said histograms.

8. (Currently Amended) The computer system of claim 7 wherein

said computing circuit multiplies frequencies of frequent values in each of said histograms to produce a estimate of join fanout of a join of said relations on said attribute.

9. (Currently Amended) The computer system of claim 8 wherein

said computing circuit multiplies a ~~number~~ frequency of a frequent value in one said histogram by an estimate of the average number infrequent values in the other histogram.

10.(Original) The computer system of claim 8 wherein

said computer system further computes a number of matching infrequent values in each said histogram by estimating a number of infrequent values in each

10/ 039,369

IBM/196

relation using said histograms, and computing from said estimates the join fanout attributable to said attribute.

11.(Currently Amended) A program product for estimating statistics on an attribute of a relation, comprising

a program of instructions executable on a computer system to form a histogram of said attribute of said relation, the histogram comprising statistics regarding tuples falling within each of a plurality of ranges of values of said attribute, said statistics being augmented to identify the most frequent values of an said attribute within at least one of said ranges, and evaluate said histogram in connection with a criterion for retrieval of data from a relation, and

a signal bearing medium bearing the program.

12.(Original) The program product of claim 11 wherein said program further comprises instructions for forming a second histogram of said attribute of a second relation, said second histogram being augmented to identify the most frequent values of said attribute, and evaluating said histograms to identify frequent values shared by said histograms.

13.(Currently Amended) The program product of claim 12, wherein said program further comprises instructions for multiplying frequencies of frequent values in each of said

10/ 039,369

IBM/196

histograms to produce a estimate of join fanout of a join of said relations on said attribute.

14.(Currently Amended) The program product of claim 13, wherein said program further comprises instructions for multiplying a number frequency of a frequent value in one said histogram by an estimate of the average number infrequent values in the other histogram.

15.(Original) The program product of claim 13 wherein said program further comprises instructions for computing a number of matching infrequent values in each said histogram by

estimating a number of infrequent values in each relation using said histograms, and
computing from said estimates the join fanout attributable to said attribute.

16.(Original) The program product of claim 11 wherein said signal bearing medium is a recordable medium.

17.(Original) The program product of claim 11 wherein said signal bearing medium is a transmission-type medium.